

# Water Quality Team Meeting Notes

Portland, Oregon  
November 8, 2004

## ***1. Greetings and Introductions.***

The November 8th meeting of the Water Quality Team, held in Portland, Oregon, was chaired by Mark Schneider of NOAA Fisheries and facilitated by Robin Harkless. Harkless led a round of introductions and a review of today's presentation. The following is a summary (not a verbatim transcript) of the items discussed and decisions made at that meeting. Please call Kathy Ceballos at 503/230-5420 with any questions about this summary.

## ***2. Credible Data/Presentation on Data Correction Software.***

You'll recall that, at the last meeting of this group, we talked a lot about data quality, validation and verification, said Jim Adams. This product was brought to my attention by Laura Hamilton, who saw a presentation from Aquatic Informatics at a conference in Vancouver and liked what she saw. We decided to invite them to give a presentation at today's meeting, Adams explained.

Various Aqua Informatics employees then led a presentation titled "Aquatic Informatics Inc. – Data Validation and Correction Tools for the Real-Time Water Quality Monitoring Industry." They touched on the following major topics:

- History of the firm – founded in 1996 in Vancouver, worked extensively in the field for Environment Canada and others, transitioned into data management, helped write BC's guidelines for validation and correction of real-time data, works with the Canadian National Research Council
- The technology trends that have been incorporated into the company's works, using scripted algorithms to provide quick and accurate throughput
- AI's customers include Environment Canada, Greenpeace, BC Hydro, a number of Native American tribes and others
- The AI scientific advisory board (list)
- Statement: "AI has the only tool today that cost-effectively manages data for the unique challenges of the real-time water quality monitoring industry."
- AI tools include: detect anomolous data, improve efficiency, validated data for further analysis, correct data using local knowledge, real-time applications or warehoused data
- AI customer benefits – reduction of errors, confidence in data, feedback into monitoring effectiveness, better information, better decisions

Peter Hudson then discussed the high-frequency hydrologic data management tool AI has developed. He touched on the following topics:

- Data management flow chart
- Raw data: outliers (graph)
- Raw data: calibration and drift (graph)
- Raw data: gaps (graph)
- Raw data: everything (graph)
- Data management flow chart: supplementary data
- Supplementary data: neighboring stations (map)
- Supplementary data: paired watersheds
- Supplementary data: meteorological data
- Supplementary data: historical data (graph)
- Supplementary data: grab samples
- Data management flow chart: modeling
- Multiple regression (sample analysis)
- Robust regression (sample analysis)
- Artificial neural networks (sample analysis)
- Auto-regressive moving average processes (sample analysis)
- Data management flow chart: data validation
- Validation statistic (conceptually) – identify some “good” “warm-up” data, calculate the expected value of the next data point etc.
- Validation statistic (mathematically) (equation) – see validation statistic white paper
- Validation statistic, equation (sample analysis)
- Data correction: model residual normalization (equation and sample analysis)
- Data correction: gap filling methodology (equation and sample analysis)
- Data correction: gap filling control (graph and sample analysis)
- Data output: validated/corrected data (list)

Is it fair to say that someone would have to be pretty familiar with this tool to use it with confidence? Sharon Churchill asked. I would say he or she would need a fair amount of local knowledge, Hudson replied. Another AI participant noted that the company is about a month away from releasing a new, more intuitive and user-friendly version of this tool for use outside the company.

Do you have a self-learning algorithm to cover a stochastic process? Churchill asked – for example, in the regulation of rivers, you need to take into account water releases, water withdrawals, meteorological processes and climate change processes – have you looked at genetic algorithms, rather than neural network algorithms, to take those factors into account? The short answer is no, Hudson replied, although we have looked at incorporating large-scale oscillations.

Another participant then described the AI roadmap, covering planned company activities through August 2005. It includes the release of beta, 1.0 and 2.0 versions of their data validation software, as well as the development of new applications, including risk assessment, trending and forecasting, as well as continuing consulting services. He noted that AI is willing to work with any potential customers, including the Corps and BPA, to incorporate any particular

features they would like to see in this tool. He added that the cost of this tool will be proportional to how much a given customer uses it.

The group also briefly discussed AI's climate change modeling efforts.

The reason I invited AI here today is that we have a huge amount of data coming into the Corps database; we're in the process of switching over to the CWMS database, said Adams. Right now, all of the data we're posting is raw data, warts and all – use it at your own risk. We're considering buying into this tool and, possibly, validating and correcting that raw data, Adams said. It might be a good idea for the WQT to designate a subgroup to begin discussing this. We may also want to explore expanding this to USGS, Reclamation and others who are collecting data, and attempting to come to some consensus in the region about how data should be treated, he said. You're looking for continuity among the various data sets? Churchill asked. that's correct, Laura Hamilton replied – it would help if there was a consensus approach to data management so that we can approach RPA 198 with some unity. If there is interest in this issue among the WQT, let's get together and talk, said Adams. If that interest is limited, the Corps will make this decision on its own.

The group briefly discussed the implications of the recently-passed Data Quality Act. In response to a request, the AI representatives agreed to email a copy of today's presentation to Schneider for posting to the WQT website. Adams added that the Corps had hoped to achieve data correction functionality prior to the next spill season, which will begin in April 2005. The real value of this tool may be to provide a much better data set to guide operations, Churchill observed. Hamilton replied that real-time operations data and archived data have been, and will continue to be, treated differently.

How do some of the other WQT participants feel about this tool? Harkless asked. Where do we need to go with this issue from here? It sounds like a good tool, but one thing I would be careful about is filling data gaps, a USGS representative observed. I think that some parameters that you measure, such as temperature, at nearby stations, could be filled in if gaps occur, added Margaret Filardo. For total dissolved gas, however, I'm not sure that would be possible, because with dissolved gas, you have independent monitoring station operations.

Have you ever had your verified data challenged in court? Churchill asked. No, the AI representatives replied.

With respect to the idea of forming a WQT subgroup to address data quality issues, said Schneider, I'm a little unsure how to proceed, because having been out of town for more than a week, I don't know the current status of the 2004 BiOp, and how this issue will be addressed. He noted that this issue will likely be addressed somewhere in the 130s, not in RPA 198. Before we decide specifically what this group is going to do, he said, let's find out what the BiOp actually says. Schneider suggested that the WQT revisit this topic at the December WQT meeting, it was so agreed.

Dave Zimmer said he would like a better understanding of what's driving the need for a

validated data set. Certainly it helps when you're doing any kind of analysis, Filardo said. Also, the federal Data Quality Act is forcing it, John Picininni observed. For the Corps, said Hamilton, we generate a total dissolved gas report every year; to produce that, we're having to make some subjective corrections to our outliers, with no scientific basis. We're keeping track of questionable data, she said, but that's hardly scientific. The other thing that's driving this is our use of the SYSTDG model; when data is bad, the model hiccoughs. We need something real-time that will tell us whether two data points are consistent with one another, said Mike Schneider.

The discussion continued in this vein for some minutes. Ultimately, Picininni distributed a recent online article describing the federal Data Quality Act. Again, it was agreed to place this topic on the agenda for the December 14 WQT meeting; without the states here at today's meeting, it's difficult to address this in detail, Mark Schneider said. The AI representatives said it would be possible for them to participate in future WQT meetings, if that would be useful. In the next month, I would suggest that a conference call with the states might be an appropriate step to bring the states up to speed, said Harkless. Adams said the Corps will be discussing the data verification issue internally, and will involve the states in those discussions. The full WQT can then discuss the possibility of the subgroup at its December meeting. In response to a question from Churchill, Adams said that, at this point, the Corps is looking only at verifying and cleaning up its own dissolved gas and temperature data – it is not looking at correcting data from the Mid-Columbia and other projects outside the federal system.

### ***3. Water Quality Data Law (Credible Data Bill).***

Mike Herold of said he has been working with the group that is constructing the mainstem Columbia TMDLs; he said Washington now has a credible data law for water quality in Washington. That requires us to create a policy on the credibility of water quality data used in the 303(d) and TMDL processes. We have to apply credible data to the TMDL development, as well as to any revisions to our water quality standards. The State of Arizona is way ahead of us; they're in the process of revising their original credible data rule. The American Farm Bureau has been trying to get state legislatures to adopt credible data laws, he said; we decided to support the bill in Washington, as long as the lobbyists and the legislature incorporated some of the changes we wanted to see. We hope the final product is useful, he said.

Herold distributed a handout briefly describing this effort, which included the web page address for the credible data effort: <http://www.ecy.wa.gov/programs/wq/qa/index.html>. This provides links to the existing QA processes Washington is using, as well as our QA policies and guidelines, the Senate bill itself, and to WDOE's quality assurance coordinator. There is a process in place for resolving disputes about water quality data QA, he said.

We're trying to work through some of the simpler policies before working our way into the meatier problems, said Herold – prioritizing basins for cleanup etc. The plan is to have an open public discussion as a part of this process. Relative to the development of your policy, is there a timeline? asked Stu McKenzie. Will you be reporting back to the WQT? How about six months from now? Herold asked. It was so agreed. He added that the overall schedule calls for

adopting the final policy by December 2006. Will this policy be integrated into the 401(k) certification process? asked Sharon Churchill. There is some linkage between the processes, was the reply. Will Ecology be following EPA's Data Quality Act guidelines in this process? asked Picininni. A lot of the existing water quality processes we already have in place are based on the EPA's guidelines, Paul Pickett replied. We'll check back in as a group six months from now, at the WQT's May meeting, said Harkless. Herold suggested that the group check the above website regularly for updates.

#### ***4. Next WQT Meeting Date.***

The next meeting of the Water Quality Team was set for Tuesday, December 14. Meeting summary prepared by Jeff Kuechle.